

HELICOPTER HUMAN FACTORS RESEARCH

David C. Nagel

Sandra G. Hart

Aerospace Human Factors Research Division

NASA-Ames Research Center

Moffett Field, CA

ABSTRACT

Helicopter flight is among the most demanding of all human-machine interactions. The inherent manual control complexities of rotorcraft are made even more challenging by the small margin for error created in certain operations, such as nap-of-the-earth (NOE) flight, by the proximity of terrain. Accident data recount numerous examples of unintended conflict between helicopters and terrain and attest to the perceptual and control difficulties associated with low altitude flight tasks.

Ames Research Center, in cooperation with the U.S. Army Aeroflightdynamics Directorate, has initiated an ambitious research program aimed at increasing safety margins for both civilian and military rotorcraft operations. The program is broad, fundamental, and focused on the development of scientific understandings and technological countermeasures. This paper reviews research being conducted in several areas: (1) workload assessment, prediction, and measure validation, (2) development of advanced displays and effective pilot/automation interfaces, (3) identification of visual cues necessary for low-level, low-visibility flight and modeling of visual flight-path control, and (4) pilot training.

INTRODUCTION

Helicopter flight is among the most demanding of all human-machine interactions. The inherent manual control complexities of rotorcraft may be made even more challenging by the small margin for error created in certain operations, such as nap-of-the-earth (NOE) flight, by the proximity of the terrain. Accident data recount numerous examples of unintended conflict between helicopters and the terrain or man-made obstacles and attest to the perceptual and control difficulties associated with low-altitude flight.

This paper reviews research being conducted in several areas. First, studies of workload are described that focus on the development and validation of various approaches to assessment and prediction. Next, we discuss the topics of displays and the development of effective pilot/automation interfaces. Since the visual sense is significantly involved in helicopter flight, particularly in the NOE environment, we describe studies that are being conducted to understand what visual cues are important and the ways in which sensor imagery and other visual aids affect pilots' perception and use of such information. Finally we review research focused on understanding flight-task decomposition and the relationship between workload and training. A companion paper in this volume by Hart, Hartzell, Voorhees, Bucher, and Shively integrates the information, understanding and technology described here into specific requirements for advanced rotorcraft development programs.

Ames Research Center, in cooperation with the U.S. Army Aeroflightdynamics Directorate, has initiated an ambitious research program aimed at increasing the margins of safety for both military and civilian helicopter operations. The program is broad, fundamental, and focused on the development of scientific understandings and technological countermeasures. The overall goal is to provide design principles, guidelines, and computational

models. Specific objectives include: (1) the design of integrated flight management displays and error-tolerant flight management systems, (2) the enhancement of visual and auditory information displays, (3) the development of valid measures and a computer-based predictive model of pilot workload for use in the design of advanced helicopters, (4) specification of human visual requirements and capabilities (in the form of a computational vision model) to serve as the basis for a machine vision system for automated NOE, and (5) guidelines for the development of training systems for advanced rotorcraft.

The primary research areas include workload assessment and prediction, pilot/automation interfaces, pilot vision research (including out-the-window visual cues, panel- and helmet-mounted displays, and biodynamic factors) and training. Figure 1 depicts a schematic overview of the program. The research efforts differ in methodology (e.g., computational modeling, empirical research, simulation and inflight testing), focus (e.g., basic or applied, vehicle-specific or relatively generic), and level of effort. An expanded Rotorcraft Human Factors Program is being formulated in response to an increasing level of interest at NASA Headquarters and urgent

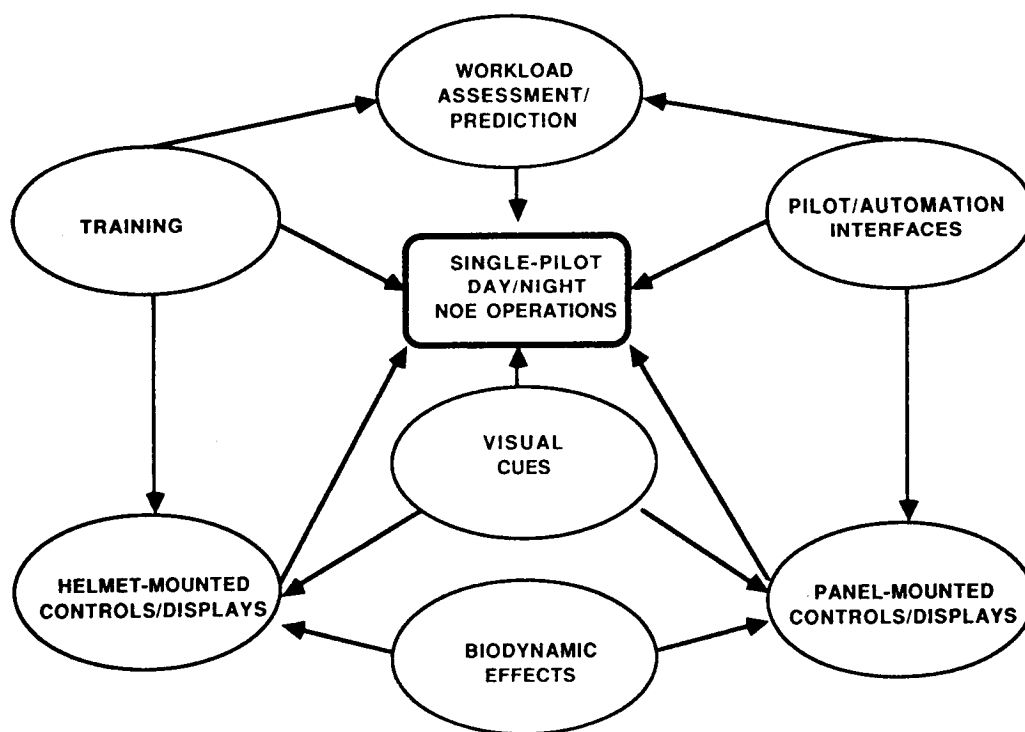


Figure 1: Helicopter Human Factors Program elements

Army research requirements prompted by the LHX program. The major challenges that will be faced by pilots flying advanced technology helicopters will be the requirement for a single pilot to conduct military and civilian missions at night and in adverse weather.

The research is conducted in-house or collaboratively with universities, industry, and government agencies. Results are transferred to user groups by establishing close ties with manufacturers, civil operators, and the military, publishing scientific research papers, participating in and sponsoring workshops and symposia, providing information, guidelines, and computer models, and contributing to the formulation of standards. In addition, the methods and theories developed by members of the scientific staff are applied to specific operational and design problems. We will summarize the accomplishments and future plans for several areas of research.

WORKLOAD RESEARCH

Introduction

The concept of workload has received increasing attention during the past decade, prompted by the realization that the human operators of advanced aircraft represent a limiting factor at the same time that their unique skills and capabilities remain an essential component. Automation has been offered as a solution to an increasing number of workload-related problems found in existing systems or that are predicted for systems under development. However, automation often simply replaces one source of workload for another, rather than accomplishing a significant reduction. In addition, there has been an ever-increasing tendency to reduce the number of crewmembers - - single-pilot operations were specified as an important goal in the early specifications for the Army's most advanced helicopter, the LHX. To achieve single-pilot capabilities, automatic subsystems must be provided to moderate the demands thus placed on the remaining crewmember. Attempts to completely replace humans by automatic systems have failed, however, because human capabilities, adaptability, and flexibility continue to surpass those of the most advanced and sophisticated systems.

If pilots could perform all of the tasks that are required of them accurately and within the allowable time constraints using available equipment, workload would be of little practical importance. Because they often cannot, accurate predictions and assessments of workload must be made during all design stages to develop optimal vehicle configurations, determine the minimum crew complement, establish mission requirements and procedures, and specify the operational envelope for specific missions and vehicles. Thus, interest in workload from an applied perspective has stemmed from the assumption that workload has a direct impact on performance and the workload imposed on the pilots is one of the final tests against which the adequacy and feasibility of operational requirements, system design, and training procedures must be tested.

It was not until ten years ago that well-controlled, theoretically-motivated research in the field of workload began to be conducted, funded by the government and industrial in-house research and development. Until very recently, however, the results of this research were not readily available to the designers and users of advanced systems because individual reports were microscopic in focus and phrased in psychological rather than engineering or aeronautical terms. Nevertheless, this research forms a scientific data base upon which meaningful, valid and reliable workload assessment tools and predictive models should be based.

In 1982, NASA formed a Workload Assessment Program to address many of the issues raised above. The goal was to merge the theoretical information about workload available from academia with the practical requirements of industrial and government organizations to develop a comprehensive definition, practically useful measures and predictors, and workload standards. Throughout the program, basic research provided answers to theoretical questions in the well-controlled environment of the laboratory while simulation and inflight research provided verification that the results were valid and meaningful in the "real world".

A Theoretical Framework

The first phase of the program was devoted to understanding the factors that influence pilot workload, evaluating existing assessment techniques, and developing new techniques. Because the workload experienced by pilots flying complex missions reflects many factors, developing a generally accepted conceptual framework within which to attack the problems of definition, measurement, and prediction has proved to be difficult; different investigators emphasize different dimensions, yet each use the same term (workload) to describe whatever it was they measured.

We defined pilot workload as the cost incurred by human operators of complex airborne systems in accomplishing the operational requirements imposed on them. It reflects the combined effects of the demands imposed by mission requirements, the information and equipment provided, the flight environment, pilots' skills and experience, the strategies they adopt, the effort they exert, and their emotional responses to the situation.

The relationships among these and other factors are depicted in Figure 2. To achieve the desired levels of overall system effectiveness, aircraft must be designed that take advantage of the capabilities of the remaining crewmembers and impose acceptable levels of workload.

The demands imposed on pilots are created by what they are asked to achieve (e.g., the objective goals of a flight and requirements for speed and precision) and the time in which they must achieve it (e.g., schedules, pro-

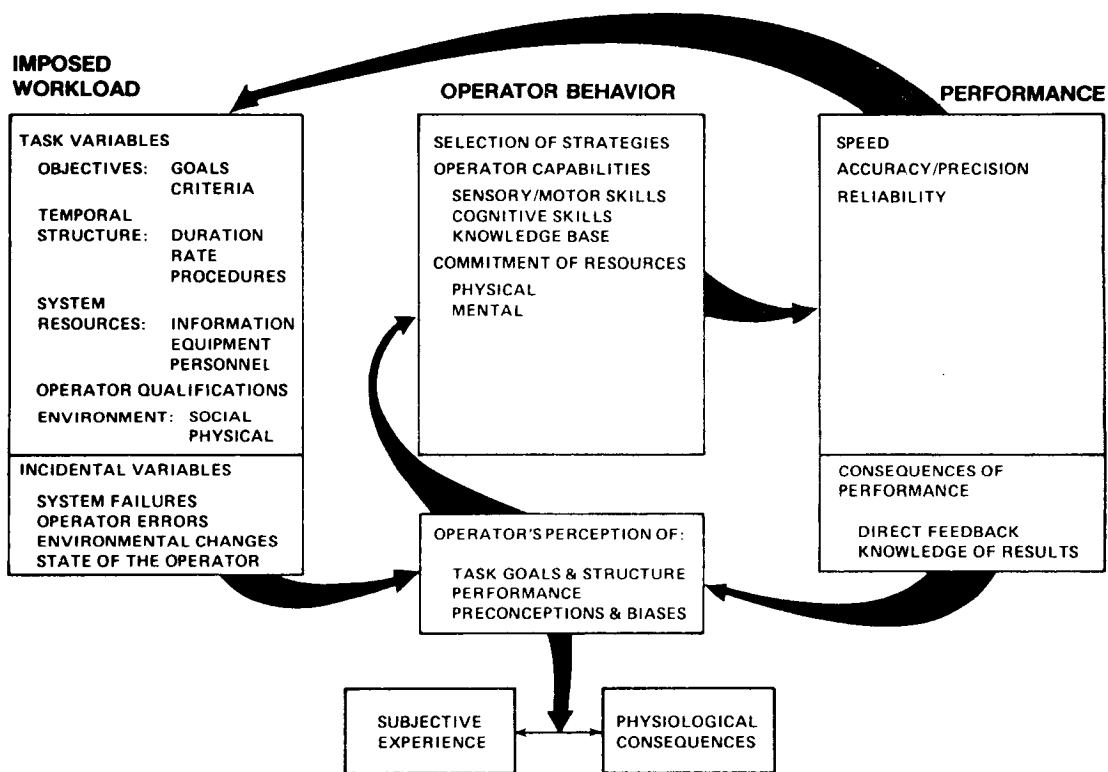


Figure 2: Conceptual framework for the analysis and prediction of workload

cedures, and deadlines). Some flight tasks are intrinsically more demanding than others, and the difficulty of almost any task can be altered by a requirement for additional speed or accuracy. System resources (e.g., controls, displays, automatic subsystems, other crewmembers, and ground support) define how pilots accomplish task demands. Poor display design, inaccessible controls, poor handling qualities, and too much or too little information can increase workload substantially. Finally, the conditions under which a task is performed (e.g., geographical location, altitude, time of day, weather) may also affect workload. For example, visual workload may be increased by low visibility, physical workload may be increased by turbulence, and threats from natural or man-made sources increase stress-related components. These elements may act independently or they may interact, enhancing or mitigating each others' effects.

Finally, the actual level of workload experienced by a particular pilot is determined by his basic skills, knowledge, and training; unskilled or inexperienced pilots experience greater workload than more skilled or experienced pilots. In addition, incorrect strategies, insufficient effort, or pilot errors can increase workload, due to the need for detecting, resolving and recovering from the problems created by the pilots themselves. Finally, pilots' expectations, previous experiences, and physical and emotional states affect their subjective experiences as well as their performance. Thus, although the "work" that is "loaded" on pilots is an important component of the workload they experience, workload may reflect a number of other factors as well.

Measures of Workload

Despite its complexity, workload is assumed to be an important and practically relevant entity and a number of valid, sensitive, and reliable measurement techniques have been developed. Given its complex nature, different measures are needed to evaluate specific components. Workload measures are usually organized into four categories: (1) objective measures of primary or secondary task performance, (2) subjective ratings, (3) physiological recordings, and (4) analytic techniques. Each type of measure has advantages and disadvantages and limitations in the range of activities and questions to which it applies; the evidence they provide may or may not be useful, depending on the situation.

Subjective Measures. Subjective ratings have been used throughout the history of workload measurement. They have face validity and are widely used and practically feasible in most environments. Furthermore, they may come closest to tapping the essence of workload because they provide a direct indication of the impact of flight-related activities on pilots that integrates the effects of many workload contributors.

One of the earliest rating techniques used in the aerospace industry was developed by pilots and engineers: the Cooper-Harper Handling Qualities Rating Scale. This scale addressed workload only indirectly, however. Other scales developed explicitly for evaluating workload were not standardized or validated and never achieved general acceptance. Furthermore, the ratings were characterized by substantial variations of opinion among raters. One of the causes of this variability was the fact that pilots responded to and considered different aspects of complex tasks when they provided ratings. In addition, the factors that contribute to workload vary between tasks. Thus, a multi-dimensional approach is better able to capture all potentially relevant factors. Research on this issue, coupled with the emerging interest in developing tools for expert elicitation by decision theorists and expert system developers, prompted the development of a rating scale that could deal with differences in the sources of workload among tasks and differences in workload definition among raters.

The NASA Task Load Index (NASA-TLX) was developed to provide an estimate of overall workload based on a weighted average of six subscales: physical demands, mental demands, time pressure, own performance, effort, and frustration. These factors represent task-related, pilot-related, and environmental factors, and are the minimum number of dimensions required to describe pilot workload experiences across different activities. The weight given to each rating reflects the importance of the factor to each rater for a specific task. This technique reflects the facts that workload experiences are created by different factors in different activities, the magnitudes of these demands vary within and between tasks, and individuals faced with apparently identical task demands experience different levels of workload. The NASA-TLX is being used extensively by government, industry, and university researchers and has been accepted as an industry standard following a recent evaluation of available measures sponsored by the FAA.

Performance Measures. Performance is the driving force behind workload evaluation in operational or manufacturing environments. It has been assumed, without empirical support, that high levels of workload will result in: (1) an increase in errors and (2) an abrupt and catastrophic decrement in performance. Instead, the typical finding is that errors occur as often when workload is too low (due to inattention) as when it is too high and that increased task demands result in strategy shifts as often as performance breakdowns.

However, measuring performance directly often provides little indication of the effort that pilots exert to achieve the obtained level of performance; as demands are increased, pilots generally put forth additional effort (to the limits of their capabilities) in order to maintain a consistent level of performance. In addition, many measures of performance reflect the characteristics of the system rather than the activities of operators directly. Finally, a common set of performance measures do not exist that can serve as workload indices across different tasks. Thus, acceptable performance-based measures of workload must reflect behavior directly and vary in response to changes in imposed task demands.

Control measures provide an objective summary of how well pilots manage an aircraft to achieve smooth and precise flight-path control. Deviations often indicate periods of time when the pilot is sufficiently overloaded by other actions that primary flight-path control suffers. In addition, the rate, content, and consequences of communications can provide an objective index of the workload imposed on pilots; a standardized taxonomy of communications has been developed in which *a priori* estimates of the workload imposed by communications tasks have been quantified. Errors and delays in response often indicate the presence of excessive workload, and the occurrence of errors is often followed by an increase in workload as the pilot attempts to resolve the consequences of the error.

Secondary Task Measures. Because performance measures do not always reflect the cost of task performance to the pilots, it has been suggested that additional tasks could be imposed that would provide an indirect indication of the resources required from a pilot to perform a primary flying task; as flight-related demands are increased, secondary task performance will degrade in direct proportion to primary task demands. The intent was to discover a secondary task "yardstick" that could be used to compare the workload of different tasks. The fact that specific secondary tasks were found to be differentially sensitive to particular types of primary tasks prompted a remarkable increase in interest in the field of workload assessment from the academic community. Competing models of attention and performance were applied to discover the structure and allocation of human resources, and a more scientific approach to the field of workload assessment evolved.

However, most secondary task measures of pilot workload are inappropriate inflight because they are difficult to implement and might compromise safety. Nevertheless, some measures, such as time estimation, can be included in the primary flight task as a natural component -- an embedded secondary task -- with minimal instrumentation and intrusion on primary task performance. This and other embedded measures have been shown to be sensitive to the workload levels encountered in different segments of flight in simulated and inflight experiments.

Physiological Measures. The earliest conceptualizations of workload focused on the physical effort required to accomplish tasks. Workload was defined in terms of physiological exertion and measures of physical effort, such as oxygen uptake and heart rate, were used to quantify this component of workload. This conceptualization of workload ignored the cognitive demands that have become an important requirement in flying advanced helicopters. Thus, measures of other physiological responses that do reflect cognitive processes (such as evoked cortical brain potentials and heart rate variability) have been investigated.

Physiological measures generally have the advantage of being unobtrusive. That is, they can be obtained without requiring attention from the pilot or interfering with a flight. In addition, since they can be recorded relatively continuously, they can reflect momentary fluctuations in workload. Finally, they provide an objective indication of involuntary physiological changes that often accompany variations in workload. The disadvantages include a lack of diagnosticity. That is, most physiological measures reflect non-specific responses to many sources of stress. These responses may reflect the demands imposed by the flight, the environment, or the pilot directly or other factors that are less directly related to workload. Cardiovascular responses do, however, provide an integrated indication of the total impact of the flight on the pilots that does not also reflect the characteristics of the system (as many performance measures do) or the pilot's biases and misconceptions (as subjective ratings do).

Heart rate reflects the stress associated with specific flight-related activities; it increases as some aspects of workload are increased. For example, heart rates are typically elevated during take-off and landing and return to baseline levels at altitude. In addition, substantially greater increases are found for the pilot-flying during take-off and landing than for the pilot-not-flying. It is possible that the feeling of responsibility and level of preparedness that must be maintained by the pilot-flying could result in their elevated levels of arousal.

Heart-rate variability is particularly sensitive to even subtle variations in mental workload; heart rate irregularity decreases as the difficulty of a task is increased. A method of obtaining an estimate of heart-rate variability on-line has been developed that provides a sensitive real-time indication of workload variations.

Inflight Verification

The final step in developing and testing these measures was inflight verification. Although simulations provide an analog of an operational environment, important elements are missing that cannot be replicated. A number of the measures developed in laboratory and simulation research were evaluated inflight in the NASA Kuiper Airborne Observatory (KAO) and in the NASA SH-3G helicopter.

In the experiment conducted in the SH-3G aircraft, evaluating the utility of different workload measures was the primary focus of the experiment. Specific missions were defined in advance and flown by each crew. The workload measures obtained for each segment were compared to predictions provided by earlier studies. In addition, portions of the flight were conducted on an instrumented flight-test range so that objective measures of performance, often unavailable inflight, could be obtained. The flight scenarios include straight and level flight above 3000 ft and contour flight, visual landings at an auxiliary site, instrument landings at airfields, hover in and out of ground effect, visual search patterns, and visual and instrument navigation conducted between Moffett Field and Crows Landing. The workload measures included pilot ratings, secondary tasks, heart rate, and selected performance indices.

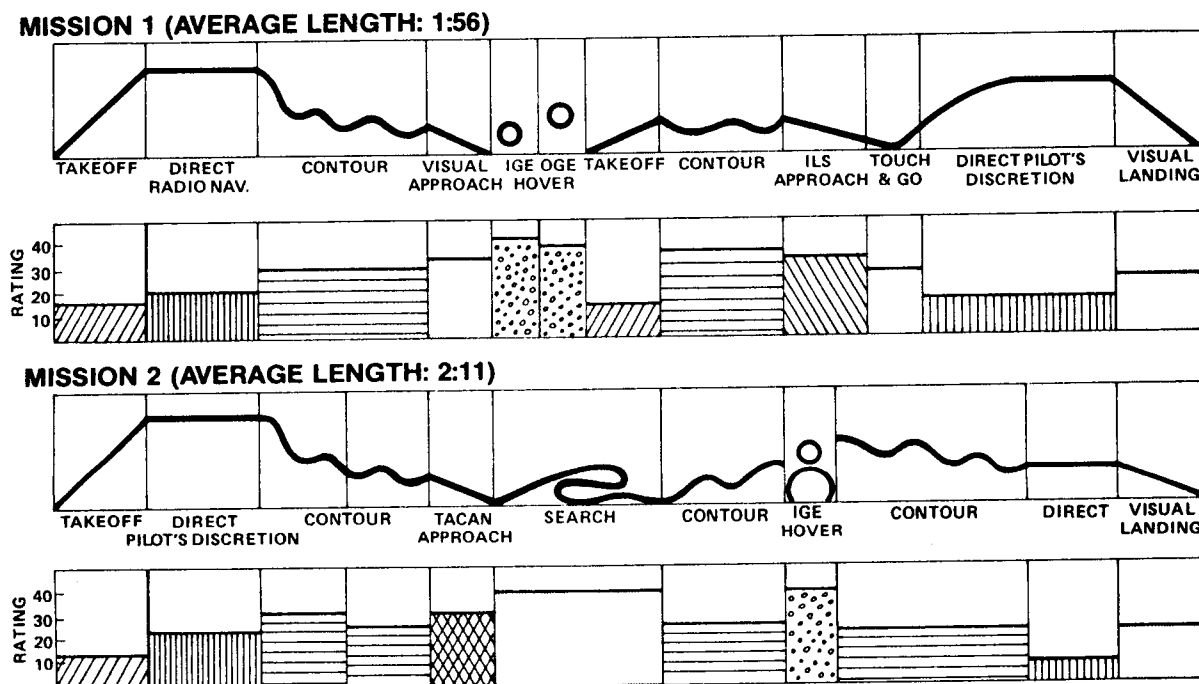


Figure 3: Flight-task segments and pilot workload ratings.

In this experiment, similar estimates of workload were obtained when the same tasks were performed at different times in the flights (Figure 3). For example, all of the visual landings were given the same, low, workload ratings. Subtle variations in tasks, however, prompted differences in workload measures that were in the predicted direction. For example, both primary and secondary performance measures and subjective ratings differed for hover tasks performed in and out of ground effect. As the environmental constraints imposed for different contour flight segments were increased, so did the measured levels of workload. As predicted, there was a high correlation between heart rate and pilot workload ratings (Figure 4). These workload-related task groupings will help in developing a taxonomy of rotorcraft tasks for a predictive workload model, and provide a

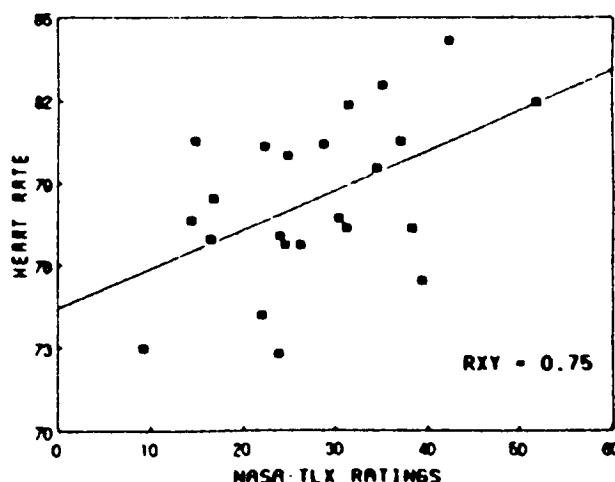


Figure 4: Relationship between heart rate and pilot ratings

data base that can be taken into the laboratory for further analysis, and can be incorporated into the Army/NASA Aircrew-Aircraft Integration (A³I) Program designers workstation.

Future Plans

The first phase of the program has been completed: the factors that contribute to pilot workload have been identified and a set of valid and practical measures have been developed. These measures are now being used to solve operational problems posed by the Army, civil operators, and industry. A micro-processor-based expert system for selecting and applying workload assessment techniques has been developed for public distribution.

The primary goals of the second phase of the workload programs are to: (1) complete and apply a computer model for workload prediction in advanced helicopters, (2) develop and publish criteria for workload (e.g., How much workload is "too much" or "too little"?), and (3) continue support for FAA, Army, civil, and industry workload efforts.

Workload Prediction

After several years of research on the structure of pilot workload and developing and applying workload assessment techniques, a computer model to predict pilot workload for current and advanced rotorcraft is being developed. In a research environment, workload predictions are essential so that known levels of workload can be imposed to evaluate candidate measures. In an applied environment, such predictions are needed to estimate the potential impact of design decisions on pilot performance early in the design process. Again, laboratory research provided the initial equations with which the workload of task elements was combined to derive predictions for complex tasks. Here, it was found that the workload of subtasks performed individually, but concurrently, could be added together to predict the performance of the combined task. Subtasks that were functionally related or shared common information, processing, or response requirements, created lower levels of workload in the combined task than would be predicted by adding the values for task components.

Experienced workload is the integrated product of many factors in addition to the objective demands that are placed on a pilot. Although workload predictions, particularly those made during the design of a new system, must necessarily focus on the objective demands that are imposed on a pilot, there are other types of information that might be included as well to enhance the predictive power of such a model. Our approach has been to start with nominal or typical flight segments or mission elements. Information about their duration, intensity, overall workload, and visual, auditory, information processing, and manual control requirements are obtained. A data base of additional tasks or events that might occur during any flight segment are identified and the same information that is obtained for the nominal segments is obtained for them. The functional relationships among specific segments and additional tasks are defined so the model can select the appropriate combination algorithms with which information about tasks and segments performed concurrently can be combined to estimate the workload of the complex task.

A preliminary model was developed based on this structure. The predictions of the model were tested in simulation research, and were highly correlated with objective and subjective measures of workload obtained in simulated flight. (Figure 5) The full model is under development. The predictions of workload made by this model will be incorporated into the A³I Human Factors Engineering Computer-Aided Designers Workstation model under development at Ames. These predictions will allow the designer of system, subsystem, or mission element for an advanced helicopter to test the effects of the design element on the potential pilot-population with a computer graphics workstation which contains many other analytic models as well. With this workstation, potential problems can be identified during the conceptual stage, thereby avoiding expensive and time-consuming cut-and-try methods.

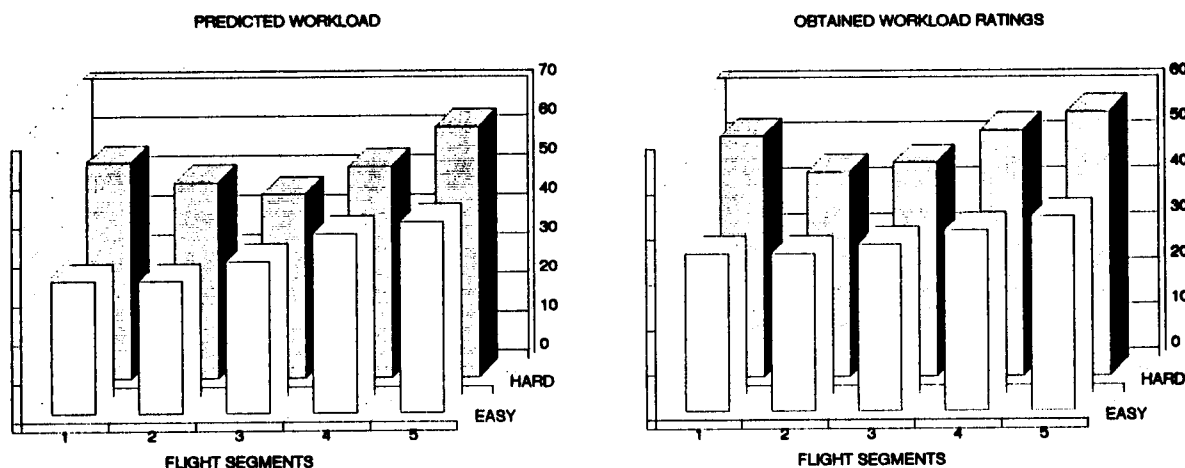


Figure 5: Obtained vs predicted workload ratings for two flight scenarios.

PILOT/VEHICLE INTERFACE

Introduction

The safest and most efficient cockpit is one in which technological advances are integrated with other cockpit systems and the pilot/vehicle interface is designed with the needs of the human user in mind so as to maintain optimal levels of pilot workload. Guidelines for implementing integrated information management, decision-aiding, and control systems are being developed. The goal is to provide specifications for systems that are tolerant of human errors and work with the pilot to accomplish mission objectives. In addition, principles based on theories of human attention, performance, communications, and learning have been used to determine the optimal interfaces for and allocation of functions between pilots and automated systems. Currently, this work is being conducted as part the Automated NOE Program, jointly with the Guidance and Control Branch. NASA researchers have tested the efficacy of some of the features that have been proposed for the next-generation of Army scout/attack helicopters (such as the LHX) to reduce the workload of a single pilot. In particular, the effects of different levels of automation and voice-interactive versus all manual input alternatives have been examined in a simulated military NOE environment.

Human-Centered Cockpit Automation

The primary goal of this program element is to integrate valid psychological principles, rules-of-thumb used in applied design, and operational training procedures to provide guidelines for the design of automation in advanced helicopters.

Rather than automating whatever functions are technically feasible, and leaving the rest to be accomplished by a pilot, cockpit automation decisions should be human-centered. Automation levels should be flexible, pilot-

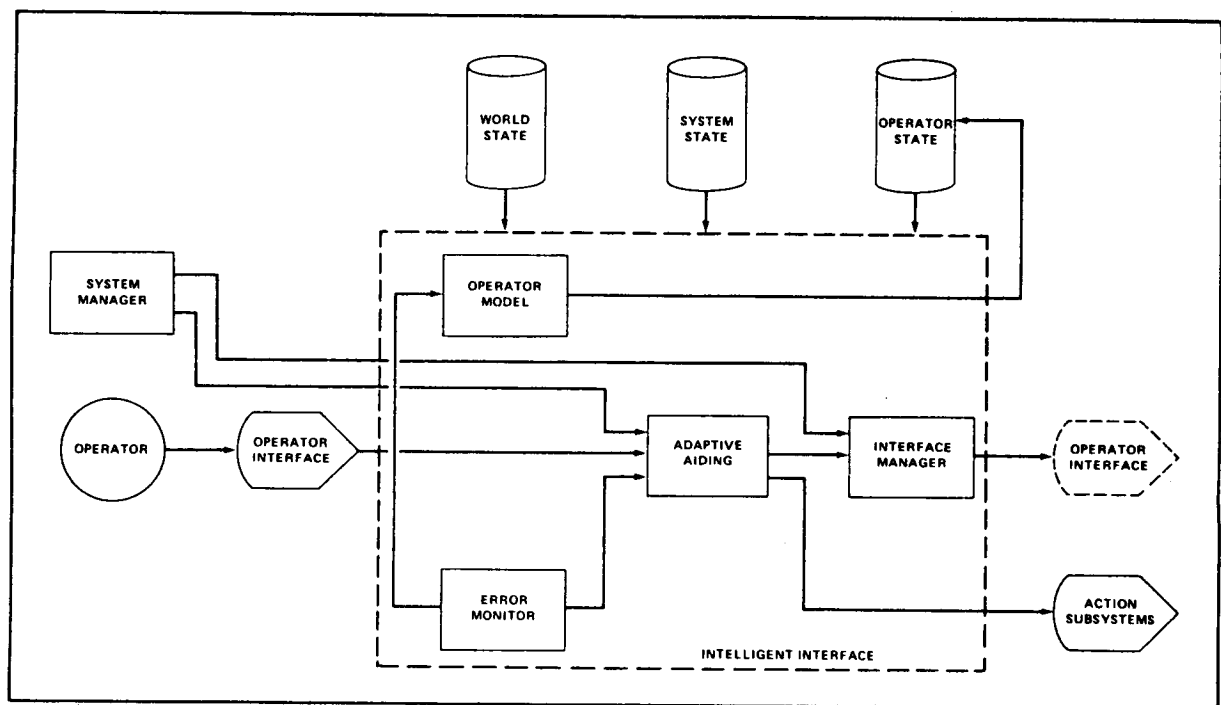


Figure 6: Architecture of an intelligent interface for a goal-directed flight management system.

selectable and driven by the intents and goals of the pilot, rather than by technology. The optimal system would be one that mimics an "ideal" copilot, providing the right level of information at the right time and assuming the level of control previously specified by the pilot at appropriate times in flight. The philosophy and conceptual design for an intelligent, goal-directed flight management system, originally developed for transport aircraft, is being investigated as the basis for the design of automated subsystems in advanced rotorcraft. (Figure 6) System objectives and pilot functions were determined in order to specify what a goal-directed flight management system should accomplish in helicopter operations and the qualitative and quantitative criteria for doing so. Refinement of this model can be obtained through basic research on human attention and performance as well as through applied research on various types of cockpit automation.

A major research effort is underway to determine the applicability of six theoretical models of human attention and performance to the design of automation for advanced rotorcraft. A complex helicopter flight simulation has been developed that mimics the information processing complexity of helicopter flight under combat conditions. Mission segments may include ingress, navigation through familiar and unfamiliar terrain, rescue, egress, and targeting. The task components include manual control, target acquisition, communications, threat response, mission management, monitoring automated subsystems, and resolution of failures. Optimal and non-optimal system configurations will be identified, based on theories of human attention, learning and performance. Model predictions are being verified with simulation research.

Two Single-Pilot Advanced Cockpit Engineering Simulations were conducted in the Ames Vertical Motion Simulator to evaluate the effectiveness of different forms of cockpit automation in reducing single-pilot workload levels to the same levels experienced by pilots in conventional two-crew configurations. Several stability and control augmentation systems, coupled with different levels of automation (e.g., hover hold, position hold, airspeed hold, altitude hold, and turn coordination) provided alone or in combination were evaluated to compare single and dual-pilot performance and workload during low-level military operations in the NOE environment. Two forms of subjective workload ratings (the Subjective Workload Assessment Technique - SWAT, and the NASA-

Task Load Index - NASA-TLX), Cooper-Harper Handling Quality ratings (HQRs), and heart rate measures were obtained to evaluate the effects of the experimental manipulations on the pilots.

A clear difference was evident between the single- and dual-pilot configurations. For example, only one configuration of the twenty tested in the first simulation was judged to be satisfactory for flying NOE with a single pilot. The workload ratings confirmed the results of the HQRs in both single- and dual-pilot configurations. In the second simulation, as in the first, it was found that different automation options were particularly beneficial for specific flight segments. For example, Figure 7 compares the results of the "best" and "worst" of the six configurations included in the second simulation. The "best" configuration (labeled 1 in Figure 7) was the basic Advanced Digital/Optical Control System (ADOCS) system with rate command, attitude stabilization, turn coordination, and heading, altitude, position, and hover hold. The "worst" system (labeled 2 in Figure 7) was an attitude command/attitude stabilization system with heading and attitude hold and turn coordination. Pilot workload and handling qualities ratings were similar during the NOE segment for all configurations, while Configuration 1 was superior for the remaining flight segments. These results, as have others, emphasize the points that automation *per se* does not necessarily have the same beneficial effects under all circumstances and that currently available forms of cockpit automation cannot reduce single-pilot workload to dual-pilot levels.

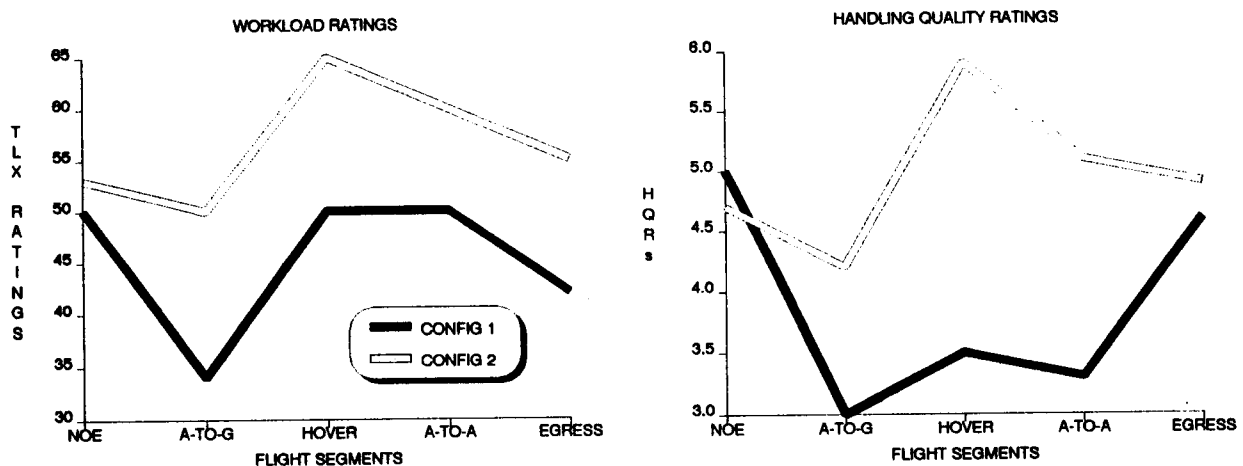


Figure 7: Comparison of workload and handling quality ratings for two stability and control augmentation systems.

Automatic Speech Recognition

Low-level rotorcraft operations are visually demanding, thus, aural displays (e.g., human, digitized, and synthesized speech messages and warning tones) have been proposed as alternative means of providing information to enhance pilots' situational awareness without further overloading their visual systems. And, since flight path control continues to impose demanding manual control requirements in rotorcraft, alternative methods for pilots to enter commands and effect subsystem selections have been investigated. For example, it has been suggested that computer-recognized vocal commands should compete less for pilots' limited resources than would manual-entered commands during activities that impose high manual control demands.

The feasibility of automatic speech recognition as a method of entering commands and information to control helicopter flight systems has been studied for many years. In general, our approach has been to utilize commercial systems, often modifying their recognition algorithms to improve recognition accuracy, and then to use the enhanced systems to develop guidelines for the application of voice recognition systems in advanced rotorcraft.

One speech recognition system was tested in a Bell Jet Ranger helicopter in level cruise, sustained turn (that created blade slap noise), approach to VNE (that created additional vibration), and hover in-ground effect (high pilot workload). Each pilot trained the system individually to his voice for the 20-word vocabulary. Recognition accuracy was not affected by cockpit noise or vibration as much as might have been expected. However, increased pilot workload did degrade recognition accuracy. The results are summarized in Table 1.

Table 1: Automatic Speech Recognition Accuracy
During Four Flight Modes

	Average Noise Level (dbA)	Recognition Accuracy (% Correct)
Level Cruise	92.5	98.3
Sustained Turn	90.0	98.1
Approach to VNE	95.0	95.8
Hover (IGE)	91.0	95.8

In a separate research effort, the multiple resources model provided predictions about tasks for which speech recognition would (and would not) be appropriate. The success of the theory in predicting performance in laboratory research prompted several simulator comparisons of speech and manual control of different subtasks in the context of advanced, high-technology helicopter missions. The task requirements included cruise, hover, and air-to-ground engagements flown by current Army Apache pilots. Some missions were flown with all-manual cockpit controls. Others were flown with a combination of manual (flight path control and weapons firing) and speech controls (weapons selection, data-burst transmission, and countermeasure activation).

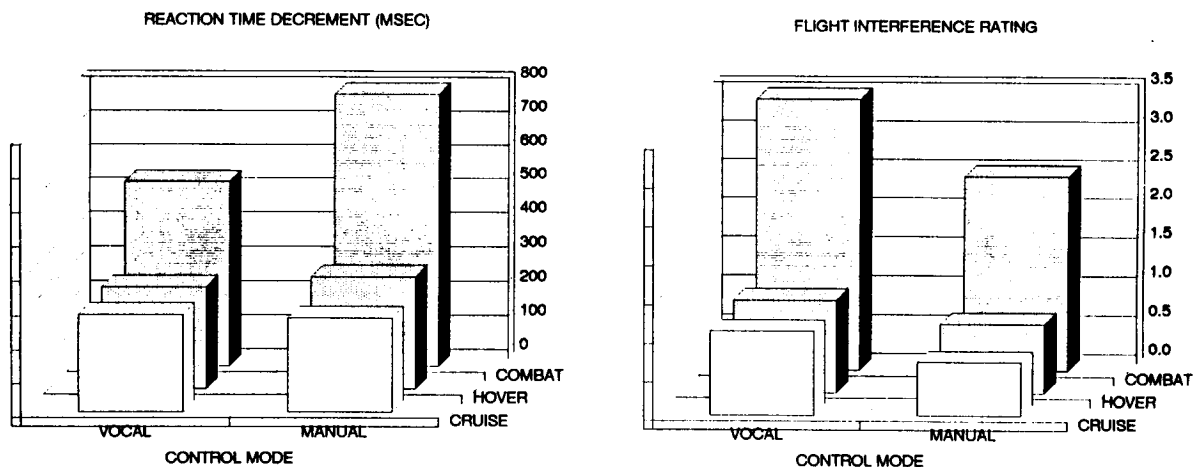


Figure 8: Comparison of speech and manual modes of entry for the selection of countermeasures

Predicted benefits (and drawbacks) of the two entry systems for different tasks (e.g., weapon selection, data burst transmission, countermeasure activation) and with different levels of automation were derived from the theoretical and laboratory research. It was expected (and found) that combat tasks, performed with no automation, would be sufficiently demanding to overload pilot's capacities for performing manual-control tasks. In this

case, an advantage for voice-interactive controls was found. However, this advantage was attenuated in the less demanding portions of the scenario. Speech controls did, in fact, improve performance by reducing competition for manual response resources. Interestingly, the performance benefit of speech controls was achieved despite pilot reports of higher workload (Figure 8). Uttering consistent voice commands to achieve adequate recognition accuracy by the voice input system demanded additional attention, as did the requirement to monitor feedback from the system (to ensure that the intended command was entered). The results of both laboratory and simulation research supported the predictions of the multiple resources theory and suggest the utility of speech controls in reducing competition for manual response resources. However, the current generation of voice recognition systems is not yet sufficiently robust or reliable to encourage their operational use for high-priority tasks. It is expected that the advantage of speech control will be further attenuated as increasing levels of automation are provided in future simulations, as they will reduce the need for continuous manual control activities. .

VISUAL RESEARCH

Introduction

The major challenge that will be faced by pilots flying advanced technology helicopters will be the requirement for a single pilot to conduct military missions at night and in adverse weather. The visual demands of low-level flight already provide a significant source of workload for two-man crews flying current-technology helicopters under visual flight conditions. Much of the information required for flight-path control at very low altitudes is obtained by monitoring the environment, with only occasional reference to flight instruments. Therefore, transition to single-pilot operations will require technological advances to facilitate and enhance pilots' abilities to extract information from the environment for immediate flight-path control and geographical orientation.

Little is known about how humans extract information about vehicle state, motion parameters, and how they orient themselves to the immediate environment from direct visual cues in good visibility. In addition, technological advances have been introduced to aid pilot performance in low visibility (such as helmet-mounted displays of sensor imagery or computer-generated symbology), although basic knowledge about human perceptual capabilities in this area is seriously deficient. Thus, research was initiated in collaboration with the Army to establish pilot performance limitations with current and advanced helmet-mounted night vision systems, specify the minimum visual cues necessary for pilots to maintain situational awareness (e.g., perceive motion, estimate range, and identify objects), and develop guidelines for the design and use of such systems based on human perceptual and performance capabilities.

The near-term goal of this program element is to specify the visual requirements in and outside of the cockpit for flight-path control, to optimize display formats and symbologies, and to augment and enhance images provided by sensors to aid in the design of adequate helmet-mounted displays for pilots to use at night and in low visibility. The long-term goal is to provide the algorithms, based on a model of the human vision system, for a machine vision system to enable automated NOE flight. This goal is a key element in the Aircraft Automation Program at Ames, and will be performed in conjunction with researchers in other divisions in the Aerospace Directorate.

Cockpit Displays for Low-Visibility Operations

There are a number of ways that a pilot's visual information requirements might be met under low-visibility conditions. Each alternative poses a number of research questions, however. Current systems for providing pilots with visual information in low-visibility conditions include light-intensifying goggles and thermal imaging systems. Computer-generated symbology or other types of graphic displays are often superimposed on a remotely-sensed video image, creating problems of display clutter, background stimulation interfering with visual tasks in the fovea, and incompatibilities when the sensor is off-axis with respect to the direction of helicopter movement.

State-of-the-art helmet-mounted displays not only limit pilots' abilities to extract motion cues (due to their narrow field of view), but also require pilots to shift their attention among sensor images and computer-generated symbology presented on the same display, panel-mounted displays, and environmental cues visible to the unaided eye. A research program has been initiated to examine the cognitive and perceptual demands thus placed on the pilots, establish guidelines about the use of monocular, binocular, and binocular HMD formats, and to propose alternative training methods to improve pilots' abilities to use such devices.

Advances in technology might enable computer-generated enhancement and display of remotely-sensed visual information. This possibility requires research on image-processing techniques, such as holography and stereoscopic systems, and matching display characteristics to the properties of the human visual system. Other systems, such as one under development at Ames might provide a low-cost "virtual" environment in which a pilot could visually explore 360 deg of computer-generated or remotely sensed environmental information with a wide-angle (120 deg) stereo display. The system might be controlled by pilot body or head position, gesture, or voice. This technology, however promising, is several years from practical application, however. Even greater advances in technology and vision modeling might eventually create machine vision systems to enable automated NOE flight.

Computational Model of Visual Flow Fields

Software tools, developed at Ames, can represent the visual information reaching the pilot's eyes, model how this information is used, and generate velocity flow fields. These analytic tools can be used to: (1) analyze the visual scene requirements (based on rate of movement, height above the ground, field of view, and terrain features) for various phases of flight, (2) determine the resolution and field of view requirements of helmet-mounted displays and obstacle-avoidance systems, (3) develop guidelines for the placement and properties of sensors to provide optimal information for human users, (4) provide guidance and control algorithms, and (5) specify the visual information requirements for NOE flight in low visibility. In addition, display formats are being developed which can provide pilots with additional information when available visual cues are inadequate (e.g., hovering with a narrow-field-of-view, monocular display).

Visual Flight-Path Control

Models of visual perception developed at Ames and aviation accident reports suggest that pilots may use inappropriate visual cues to control altitude. For example, a model of human slant perception was developed that can predict the errors that will occur under conditions of low visibility or on relatively featureless terrain. Even for landings that occur at airfields, the absence of visually perceptible perspective lines converging on the targeted touchdown point early in the approach can contribute to "landing short" accidents. The problem is particularly acute for the steep angles of approach that helicopters are capable of flying. Additional research is being conducted to determine the information that is provided by motion cues (such as naturally occurring optical flow patterns or those simulated on a cockpit display) to extend the utility of the model.

A part-task simulation was conducted to determine how pilots extract and use vehicle motion cues (represented by stylized perspective displays such as depicted in Figure 9a). Horizontal perspective lines (orthogonal to the direction of flight) contributed to better altitude control than did perspective lines (parallel to the direction of flight). Optical motion associated with lateral translations was misinterpreted as altitude changes, further contributing to altitude-control errors (Figure 9b).

Additional research was conducted to examine how pilots extract cues from panel- and helmet-mounted displays to maintain visual control of altitude while tracking targets moving off-axis to the direction of flight. The performance envelopes for three-dimensional target tracking using visually-coupled devices was examined with impoverished (perspective grid displays) and high-fidelity (computer-generated terrain) visual scenes. A part-task simulation established range, azimuth, and elevation parameters. Optical velocities accounted for most of the tracking errors; velocities as low as 30-40 deg/sec resulted in degraded performance. As the slant range of a target was decreased, a constant ground error was represented by an increase in displayed error on the helmet-

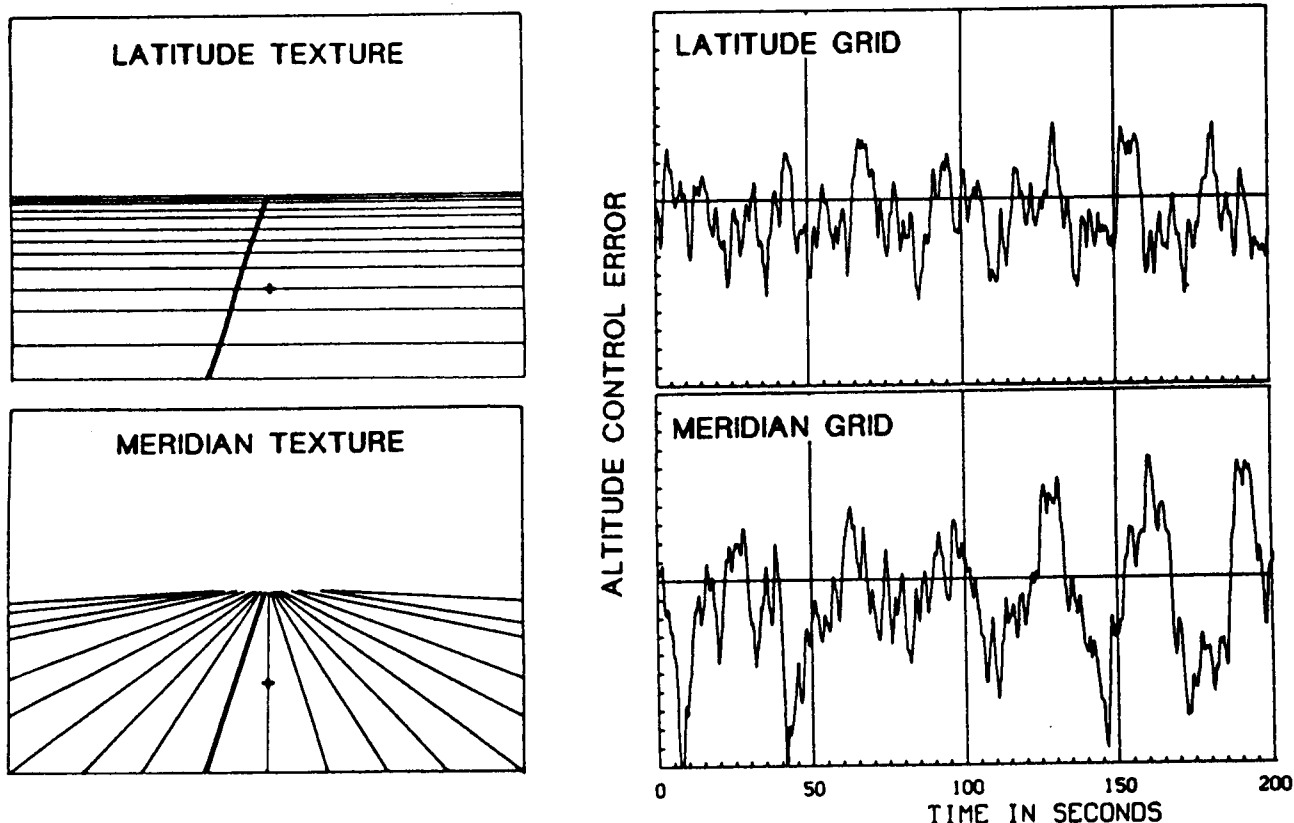


Figure 9: Altitude control performance using texture-augmented displays.

mounted display. Thus, line-of-sight errors projected onto the ground were minimized, rather than display error (the basic geometry is depicted in Figure 10).

A high-fidelity helicopter simulation evaluated the performance consequences of off-axis tracking in a more realistic setting. Army Apache pilots tracked moving targets from a hover or while in motion using information displayed on a Honeywell IHADSS. Target tracking was accomplished by the pilot-flying (Manual) or the pilot-not-flying (Automatic) under simulated day and night conditions. Preliminary analyses indicate that pilots had difficulty conducting simultaneous target tracking and aircraft control tasks. Again, optical velocity accounted for most of the tracking errors, ground errors remained relatively constant (although displayed error increased) as slant angle increased (Figure 11), and tracking performance was worse with curved flight paths than with straight.

Biodynamic Factors

The detrimental effects of vibration on visual acuity have been well documented for direct vision. They are even more severe with helmet-mounted displays where decrements are caused by relative motion between a displayed image (due to involuntary, vibration-induced head motions) and the eye. The normal vestibular-ocular reflex induces eye movements that oppose those of the head to maintain a stationary point of regard. While appropriate for viewing panel-mounted displays, it is not appropriate for helmet-mounted displays; relative motion is produced between the image on the head-coupled display and the eye, resulting in retinal blurring, increased errors and longer responses.

Based on a computer simulation of the vibration frequencies of helicopters, an adaptive noise-canceling technique has been developed at Ames, that minimizes the relative motion between viewed images and the eye by shifting displayed images in the same direction and magnitude as the induced reflexive eye movement. This filter stabilizes the images in space while still allowing the low-frequency voluntary head motions that are required for aiming accuracy.

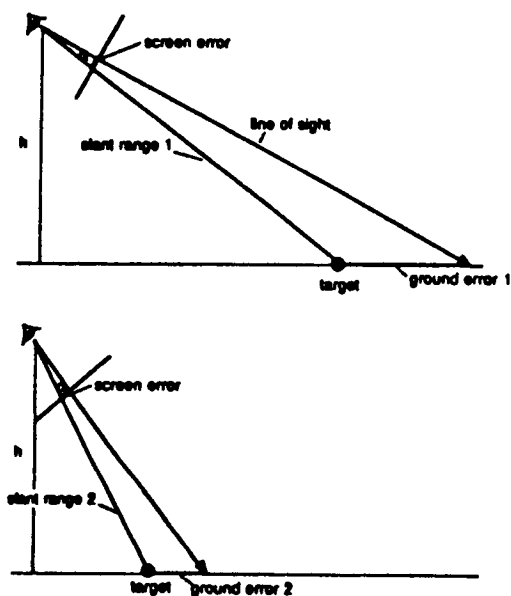


Figure 10: Screen error vs ground error: Basic Geometry

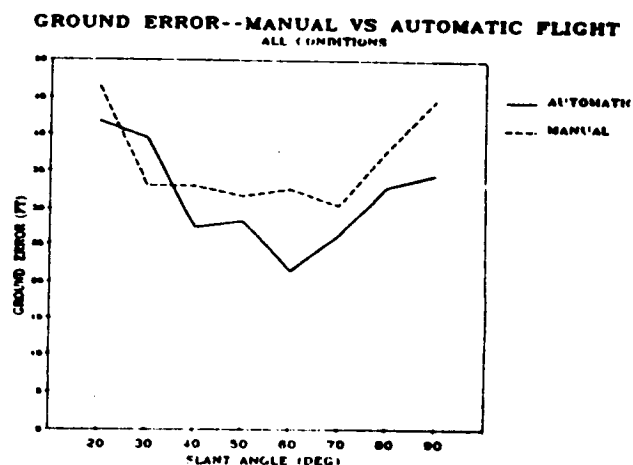
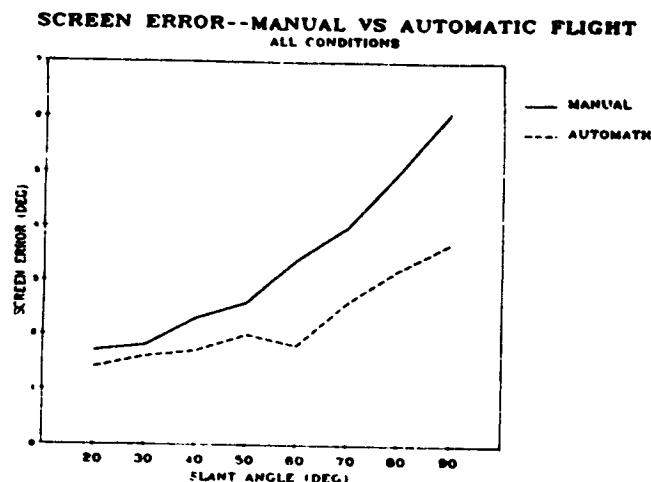


Figure 11: Manual vs automatic flight: Obtained screen error vs ground error.

Display Resolution

The problem of display resolution has been raised, particularly in conjunction with helmet-mounted pilot night vision systems. Research is required to determine whether the field of view is the key to enhancing the effectiveness of a night vision system, or rather the display resolution of the images, the addition of binocular cues, or other factors.

The computational model of human vision developed at Ames can be used to specify optimal spatial and temporal sampling procedures to enhance human visual performance using computer-generated displays, video displays provided by remote sensors, or computer-generated symbology superimposed on video displays. Depending on the display resolution available, accurate estimates can be made about visual information frequencies that can be displayed without distortion. In addition, by emulating the irregular sampling characteristics of the human vision system, the problem of aliasing can be avoided when generating visual displays.

Expert System for Symbology Evaluation

Most advanced helmet-mounted display designs include superimposed, computer-generated instrument displays. Although it is often possible to evaluate candidate symbologies empirically, a more effective method is to evaluate them in software. For example, an expert system has been developed by the Perception and Cognition

Group at Ames to automate the evaluation of helicopter display symbologies. The Ames Vision Model was used to compute the perceptual distances among alternative symbols and fonts to provide objective criteria for selecting perceptually distinct symbols.

TRAINING

The training research element includes theoretical development of optimal training strategies, application of these strategies in operational environments, and formulation of evaluation criteria for training programs. Our interest in training evolved from its apparent relationship with workload. Training is often proposed as a solution to workload problems, as it is assumed that both performance and workload are equally improved by training. However, the two research areas rarely, if ever, overlap and there is little empirical evidence to support such assumptions. In addition, the workload encountered by trainees must be monitored to ensure that: it is low enough to allow learning to take place, (2) training elements and promotion rules are derived logically, and (3) training time is optimized. Since the cost of training is escalating rapidly, low-cost, effective training methods are urgently needed.

NASA/Army Workshops

Ames co-sponsored two workshops with the U. S. Army to initiate this program element. The topic of the first workshop was the relationship between workload and training. The topic of the second workshop was individual differences in pilot selection, training, workload, and operational performance. Participants were invited from academia, industry and the government to discuss workload and training and their relationships in the context of advanced helicopter operations. The results of the workshop have been summarized in an Executive Summary and will be published in book form. The meeting was a great success in acquainting members of different research communities, revealing their problems, and discussing why there is such a limited flow of information and support among industry, academic and government research laboratories.

Theoretical Development

As a consequence of these two workshops, there has been a significant increase in research directed toward understanding the relationships among workload, training, and performance and in developing efficient and effective training strategies. Although training is essential in achieving acceptable levels of performance for complex systems, and it is often proposed as a solution for workload-related problems, the assumption that both performance and workload are improved equally by training appears to be unfounded. Furthermore, trainee workload must be monitored to ensure that it is low enough (for learning to take place), yet high enough to make optimal use of training resources.

Part-task training is one method by which acceptable levels of trainee workload can be achieved during the initial exposure to complex and difficult tasks. In part-task training, specific task elements are singled out for individual practice. The goal is to reduce the workload of these tasks (by achieving automatic performance) so they will not consume the cognitive resources required to learn the remainder of the task. A disadvantage is that these tasks must be re-integrated into the whole task to ensure that they can be performed collectively.

An alternative method has been suggested in which the whole task forms the basis of training, but priority instructions focus the trainee's attention on one element of the task at a time. This form of embedded part-task training achieves acceptable levels of workload by requiring that trainees learn and attend to only one element of the task at a time. Since the component tasks are presented in the context of the entire task, however, some time-sharing skills are developed during initial training, and integrating the component skills to promote efficient overall task performance presents less of a problem.

These and other training approaches have been evaluated in laboratory research and in the context of helicopter flight training. The goal is to identify which skills and task components are most likely to benefit from specific training strategies and to develop guidelines for developing training programs. Existing theoretical predictions form the basis of this research, although their applicability to complex skills, such as helicopter flight, had to be established. For example, laboratory research has suggested that performance and workload improvements would be found for "consistent" task components but not for "inconsistent" components. A consistent component is one

for which specific information, display configurations, or system states are invariably associated with a specific response. By contrast, inconsistent components may require different responses each time they occur, and neither performance nor workload improvements occur as training progresses. This theory would predict that pilots would benefit from part-task training on target acquisition (a "consistent" task) but not from part-task training on vehicle-control training (a less consistent task). Alternatively, for tasks in which time-sharing is the most critical required skill, other theories would predict that part-task training methods would be less effective than whole-task training, because learning how to integrate or time-share two tasks that require common visual/manual resources can only be learned in the context of the whole task. To conduct

this research element, an efficient method of performing task and skill analyses was developed to identify the specific skills that might be improved by either part-task or whole-task training methods. Several alternative training methods, including part-task, whole-task, and priority manipulation within a whole-task, have been examined in simulation research. For example, in a recent helicopter simulation, it was found that part-task training on more consistent task components (e.g., target acquisition) resulted in better vehicle-control performance in the integrated task than did part-task training on vehicle control *per se* (Figure 12). However, neither part-task training method was as effective as whole task training, as this whole-task training allowed pilots sufficient opportunity to develop necessary time-sharing skills in context. The results of this and other simulation research emphasize the importance of matching training procedures to the characteristics of individual task elements as well as to the structure of the integrated, target activity to achieve the most effective and efficient use of training time and facilities.

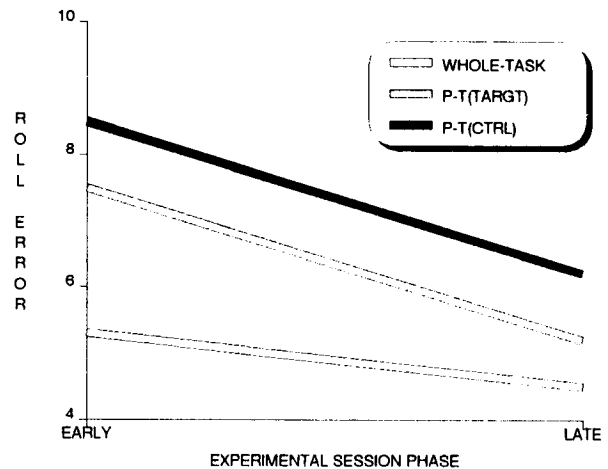


Figure 12: Comparison of three training methods in the

Training Procedures for Advanced Night-Vision Systems

Training pilots to use helmet-mounted, sensor-driven, night vision systems is one of the most demanding challenges faced by military and civilian users. Although these systems allow pilots to perform low-level missions at night and in low-visibility, they impose high workload and pilots are generally not able to achieve the same performance levels possible with direct vision. Furthermore, the cost of training is extremely high; training is currently conducted entirely in flight as adequate simulations do not yet exist. Thus, providing initial training in micro-processor-based, part-task simulators might provide an environment in which the perceptual and cognitive problems characteristic of this system could be resolved prior to the beginning of inflight training.

Pilots experience considerable difficulty during initial training with helmet-mounted pilot night vision systems. In addition, some of these difficulties persist even after many hours of use and skills degrade quickly if they are not used. A survey was conducted in which several key problem areas emerged: low resolution, reduced field of view, inadequate image definition, monocular display format, offset sensor location, superimposed HUD symbology, and helmet fit and weight. These system-related problems result in difficulties in detecting, identifying, and tracking objects, binocular rivalry (between the aided and unaided eyes), motion parallax, poor motion, depth, and range perceptions, and visual fatigue. These, in turn, limit performance and mission duration.

These results prompted research to identify the underlying perceptual and information processing mechanisms which must be addressed during training and to establish optimal learning strategies. The goal is to develop guidelines for a low-cost, microprocessor-based training program which will allow pilots to develop the skills necessary to use such systems effectively.

In parallel, a research program was established to evaluate the effectiveness of microprocessor-based training programs in the acquisition and performance of general flight-management skills. The goal is to evaluate the effectiveness of complex computer games in improving the acquisition and performance of flight management skills. The underlying rationale is that the strategies and cognitive skills learned by interacting with low-cost computer games can be generalized to the learning and performance of complex operational tasks. Cognitive skills training will be provided to half of the members of an incoming class of pilots during basic flight training. The total training times, flight school retention statistics, and flying performance of the two groups of pilots will be compared and a longitudinal study of their operational performance will be conducted.

To date, a task/skill analysis has been completed and the training program has been developed. Field tests began in October 1987 with a group of incoming pilot-trainees. Total training times, flight school retention statistics, and flying performance will be compared to those of pilots in the standard training program. If the cognitive-skills-training procedures prove to be effective, then a ground-based training system will be developed to improve pilot performance with night vision systems. This system will train pilots how to interpret superimposed symbology, focus and shift their visual attention, compensate for the loss of peripheral cues, and adapt to the distortions and resolution limitations of current (and projected) helmet-mounted, night vision systems.

SUMMARY

At each stage in the research process, information obtained in more realistic situations was used to refine theoretical models and provide the focus for well-controlled laboratory studies to address specific issues. By moving back and forth between the three research environments, the requirements of theoretical development can be balanced against the requirements of the "real world". Furthermore, operational relevance can be insured at the same time that the predictive advantages of a theoretical foundation can be maintained. If theoretical researchers cannot become familiar with applied problems (through participation in simulation and inflight research) and if designers, engineers, and operational test and evaluation personnel are not exposed to the advantages of experimental control, a theoretical foundation, and the use of validated measures, the advances made in none of these fields can be capitalized upon.

This is the unique role that a government research laboratory, such as Ames Research Center, can play. Here, scientists, engineers, and pilots work in close proximity to each other and can conduct collaborative research in which each can take advantage of the other's knowledge and tools to provide a strong research foundation that can be transferred to industry. If nothing else, this environment creates a unique opportunity for each group to learn the other's language. This provides a vehicle for translating the considerable data base available in academia and the pragmatic experiences of designers, engineers, and pilots into a useful body of knowledge.